

*Croatian Journal of Philosophy*  
Vol. XVIII, No. 53, 2018

# *Intuitions: Epistemology and Metaphysics of Language*

NENAD MIŠČEVIĆ  
*University of Maribor, Maribor, Slovenia*

*The paper addresses the issues about grammatical intuitions in a programmatic sketch. The first part deals with epistemology of such intuitions and defends a moderate Voice-of-competence view in discussion with Michael Devitt, the ordinarist, who sees them as products of general intelligence or Central Processing Unit. The second part deals with the problem for their validity and offers a compromise solution: linguistic intuitions are valid because their object the standard linguistic entities, are production -and response-dependent. Competence does dictate what is correct, and what is not, the order of determination goes from the internal to the external, or external-seeming language items. An external token string has linguistic properties because it would be interpreted as having them by the normal language-hearer and would be produced by a process that would form it respecting the nature of these properties. The solution is briefly situated on the map of general response-dependence.*

**Keywords:** Intuition, competence, ordinarism, response-dependence.

## 1. Introduction

The paper is dedicated to Michael Devitt (to be called just “Michael” in the sequel), and continues our long and fruitful discussion about linguistic, in particular syntactic, intuitions. It has two parts, the first more epistemological, but on the non-normative side, the second more metaphysical. Let me say a few words about each. The main epistemological debate in philosophy of language concerns people’s linguistic intuitions. Let me borrow an example from Isaac (2008: 178) to illustrate the kind of items that will be discussed in the paper. Suppose a linguist confronts John, a native speaker, with a following sentence:

(H\*) “Herself loves Mary”.

Is this a sentence of your language, the linguist asks. No, answers John. John has clear feeling that something is amiss with (H\*), and the feeling gives rise to the judgment and belief expressed by his report “No”. The feeling, the belief-state and the judgment, and sometimes even the report (all of them, or at least some of them) are called “intuition”; I will reserve the term for the first three of them. A famous recent tradition in philosophy of language and linguistics, initiated by Chomsky takes intuitions as the main source of data for the linguist. The tradition equally sees them as products of linguistic competence; the term is ambiguous between the ability and the mechanism, so I will use it for both, but most often for the later. I agree with the general idea, but I would like to allow that many opinions that people voice as their “intuitions” contain a lot of material not produced by, and therefore not revelatory of, the pure linguistic competence. (I have been defending similar views about other competences like the, logical and the spatial-geometrical ones). I have been calling my own line “the Moderate Voice-of-competence view”, since it takes competence as basic, but allows for very strong interferences or external contributions to the production of intuition. John’s intuition at best points to a discreet voice of competence. So, the Moderate Voice-of-Competence view that I would like to develop and briefly defend here, claims that intuitions form a kind, albeit relatively superficial one, shearing their phenomenal properties classically described by Descartes in his *Regulae* and *Principles*. in terms of “clear and distinct cognition”, of being „present and apparent to an attentive mind, (Principle XLV, 168). And there is a capacity, or rather several of them associated with intuition. Further, they are extroverted, turned towards the items they are explicitly about, and normatively answerable to them, since they want them to teach us about things “outside” (possibly in Platonic heaven), not merely about our representation(s) of them). The view also takes seriously the actual dialectics of having intuitions: asking (or being asked) a question, then going through imagining a scenario, if necessary putting oneself in the shoes of the imagined person (or a person-like entity, say, zombie), and then giving a simple, preliminary answer to the question, formulating the immediate intuition, often to be developed by considering other examples, and so on. This intuition-related effort involves a lot more than mere inference following rules of logic. Further, the view is for the most part committed to realism about the objects of intuitions, and in the second part of the paper I develop a moderately realistic view about linguistic entities. The view is very keen on the explainability of intuitions. Finally, it offers a somewhat complex answer about their normative epistemic status, tilted towards aposteriority: although intuitions are *prima facie a priori*, their reflective justification has a rich structure in which a posteriori elements play a crucial role.

The theories on the opposite end favors deny the specific nature of intuitions, insist on holistic non-*a priori* justification, and see them as

products of general intelligence. Let me call this pole “ordinarism”; it has been invented long time ago in ethics, with Ewing (1971) and other moral intuitionists, like More (1991), in some of their moods), but is being now re-discovered for other domains, generalized and put forward in vigorous manner by Michael (Devitt 2005, 2006, 2012), and specifically for philosophical intuitions T. Williamson (2007).

I disagree with Michael’s mistrust. He is an empiricist about linguistic intuitions, and he is comparing them to expert judgments, paleontologist in the field searching for fossils. She sees a bit of white stone sticking through grey rock, and responds immediately “a pig’s jawbone.” (Devitt 2012: 560), art experts correctly judging an allegedly sixth-century Greek marble statue to be a fake; of the tennis coach, Vic Braden, correctly judging a serve to be a fault before the ball hits the ground (Devitt 2006a: 104).

The second part of the paper deals with the further problem: where does the validity of our intuitions come from? In order to answer it, it briefly visits the location problem for language, whether it is essentially E-language situated in the outside world, or I-language, inhabiting only the mind of speaker-hearer) and offers a compromise solution: standard linguistic entities are production- and response-dependent. An external token string has linguistic properties because it would be interpreted as having them by the normal language-hearer, and would be produced by a process that would form it respecting the nature of these properties. The solution is then briefly situated on the map of general response-dependence.

## 2. *Linguistic intuitions: The voice of competence*

### 2.1. *Intuitionism-competentialism*

Let us then start from scratch. The typical context in which linguists speak about intuitions is the one of linguistic research. The linguist presents a string of sounds (phonemes, letters) to the native speaker of the language investigated (often to oneself, if the language is one’s mother tongue), and asks her to decide if this would be a sentence of her language. The immediate judgments prompted by the question are described as linguistic intuitions.

I shall here re-use my older examples, since I have changed my explanation at one important point, thanks to Michael, and I want to stress the continuity and discontinuity with my former presentations. To have a handy example I have borrowed from a fine introduction to Chomsky by John Collins (2008) a pair of sentences testing the predicate nominal agreement:

(W) They want to be teachers.

(W\*) \*They want to be teacher.

Imagine a native speaker, Ann, accepting the first and rejecting the second. Perhaps Ann sort of rehearsed the sentences in her inner fore,

asking herself whether she would say them, simulating actual saying, as the standard description goes. The next example is our already mentioned

(H\*) “Herself loves Mary”.

rejected in our story by the naïve subject John. Finally, here is the co-reference example.

(M) Mary knows that Jane loves herself.

(M) can be taken in two ways: the incorrect one, according to which Jane loves Mary, and the correct one, which is obvious. So, suppose the linguist asks the naïve subject John two questions and receives the following answers:

Q: Does Mary know that Jane loves her, Mary?

A: Of course, not.

Q: So, whom does Jane love?

A: She loves herself, Jane.

So, how are Ann’s and John’s cognitive apparatuses arriving to the verdict? Following the lead from the Chomskyan tradition, I would claim that it is mobilizing the particular competence, i.e. the same cognitive resource that produces or fails to produce similar sentences in real-life speaking. It is the competence itself that is doing the work, the central processor at best just passively reports the verdict of the competence, which is the intuition. Michael would claim that Ann’s and John’s apparatuses are mobilizing the cognitive resource that is normally in charge of understanding sentences. Ann’s resource outputs the verdict Yes for the first sentence (W), and No, for the second (W\*), in some neural code. John’s outputs a No for the (H\*) sentence. If the resource is competence (the particular, i.e. linguistic one, then the result is the voice of competence.

I will take it as agreed by all sides that the first stage must be a tentative production of the sentence, and I would add that it is being rehearsed and analyzed by Ann’s competence. I mentioned that competence presumably comes out with some kind of answer, some Yes or No signal. I will argue that this is the most important element, the core, of the final intuition. The next stage is empirical theorizing at sub-personal level; Ann’s central processor, CP for short, has to interpret the message, decides how to treat it, and then translate the message into the spontaneous belief, what we call intuition. The rest is reporting, producing the verbal output. This linguistic intuitional output has a very narrow range. Intuitions reported are formulated in an austere vocabulary, featuring mainly “acceptable” vs. “non-acceptable”. So what does this tell us about folk-concept of grammaticality? Distinguish an egocentric minimal concept, expressed by “I wouldn’t say S”/ “I would say S” (for some target sentence S), from rich socio-centric concept IS A PART OF MY COMMUNITY LANGUAGE, predicated of “S”. Two related points: first, for producing relevant intuition, only the ego-

centric minimal concept is needed. Second, this concept is not clearly empirical, culled from past experience with one’s own sayings. Its application is typically guided by immediate promptings of competence: if you ask me about a sentence in my mother tongue, Croatian, what I primarily do is that I either just “hear” that this is not what I would say, or I try to produce the sentence internally. My verdict is then an immediate Yes (or No, or, in the worst case Yes-and-No, e.g. if I am very drunk). And the basic data for the linguist are that Ann would say such-an-such and would mean such-and-such by the given expression. The rest is sociological theory. Remember, when a Chomskyan speaks about one’s language, he means one’s idiolect. Ann’s opinion whether she is a typical representative of a community, and whether her English is good English are beside the point. Does Ann have to think that she is competent in the very language (idiolect) she is speaking? The very question sounds ludicrous. Why is this relevant? Because stating the judgment about a sentence in one’s idiolect requires far less theory, if any at all, than reflection about social usage.

Let me rerun a simple scheme of the hypothetical production of a syntactic intuition I proposed a decade and half ago (Mišćević 2016), with an important correction inspired by Michael’s criticism. Here are two sentences

- (W) They want to be teachers.
- (W\*) \*They want to be teacher.

Imagine a native speaker, Ann, accepting the first and rejecting the second. Ann can rehearse the sentences in her inner fore, simulating producing it, or her linguistic cognitive apparatus can just analyze the sentence heard. Or, she can go on asking herself whether she would say them, simulating actual saying, as the standard description goes.

1	2	3	4
<i>Processing the target sentence (heard or simulated)</i>	<i>immediate, spontaneous verdict by specialized competence (intuition core)</i>	<i>empirical testing at sub-personal level</i>	<i>intuition</i>

In my earlier presentations of the MoVoC theory I have placed in the first box only simulation; Michael has kindly pointed out to my mistake, so I have enriched the content of the first box in in the meantime. The immediate spontaneous answer is the datum used by the central processor to arrive at the belief state, intuition proper. The generation of linguistic intuition-states seems to be rather isolated, independent of general intelligence, employed in stage 3 theorizing. of our flow-chart. “CP” stands for “central processor”, the general intelligence.

Similarly, for more complicated structures, like the ones involving co-reference.

(M) Mary knows that Jane loves herself.

The immediate spontaneous answer is the datum used by the central processor to arrive at the belief state, intuition proper. The generation of linguistic intuition-states seems to be rather isolated, independent of general intelligence, employed in stage 3 theorizing. of our flow-chart. Second, it seems that the immediate, spontaneous answer (the item no. 2 on the flow chart) , is the item that deserves to be called “intuition”. It is not necessary that verdict is unconscious; and if and when it is conscious, it seems to be the prime, if not the unique candidate for immediate (and perhaps obvious and compelling) judgment. In that case, the empirical testing proposal is just a verbal maneuver; calling the last stage, no. 4 “intuition”, instead of stage no. 2.

On the view I propose, “the Moderate Voice-of-competence view”, the answers in no. 2, no. 4. and indirectly in no. 5 are often produced by, and in such a case, revelatory of the linguistic competence. Only the competence has access to grammar, whatever its nature.

The ordinarist theories all this, since they want to deny the specific nature of intuitions. The ordinarist’s flow-chart has roughly the following shape:

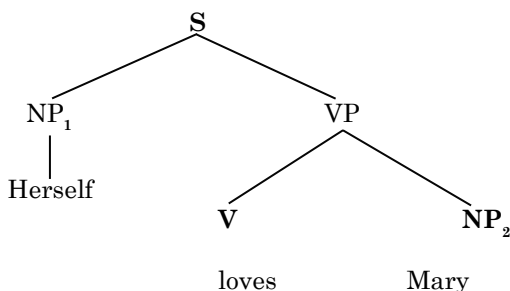
1	2	3
<i>Processing the target sentence (heard or simulated)</i>	<i>empirical testing by central processor</i>	<i>intuition (so-called)</i>

The crucial question is whether and how the central processor, i.e. general intelligence accessing memory, can quickly analyze complicated syntactic structures involved in ordinary sentences, like the ones involving co-reference. The empiricist-ordinarist answer is that the capacity is due to empirical exercise, like the one characterizing the expert knowledge of, say, a paleontologist, or a good tennis player. (I call it ordinarist, since it fits nicely with other views that see intuitions as items of ordinary knowledge, most famous of which is the view of Tim Williamson (2008). Linguistic ordinarism denies that seemingly linguistic judgments form a significant epistemic kind, intuitions, and that there is a distinct capacity producing them.

Roughly, the ordinarist hopes that the contribution of the competence is minimal, and the holistic contribution of CP maximal and essential. For her, intuitions are basically the products of holistic theorizing, not of special, dedicated competence. This is why they are not special, why they contain so much empirical material, and why it is wrong to take them to be a priori. I find it incredible. Nothing in our ordinary empirical knowledge points to a general structure-recognizing empirical ability of this power. And, in his forthcoming book Devitt (2019) is quite skeptical about people’s semantic intuitions! He is not

rejecting them altogether, but does not want to use them as evidence, and criticizes others for using them exclusively as evidence. One can read him in a more radical and in a more moderate way: on the radical reading, he is rejecting them altogether, on the moderate reading he is just demanding the theorist to re-check her intuition appealing to non-intuitional sources, like elicited production or analysis of the corpus. In the discussion, Michael opted for the moderate reading, but some of his formulation suggest the radical version. I will say more about the variants of the ordinalist view in the sequel. In order to determine who is right, we have to discuss the proposals following the flow-chart(s) stage by stage.

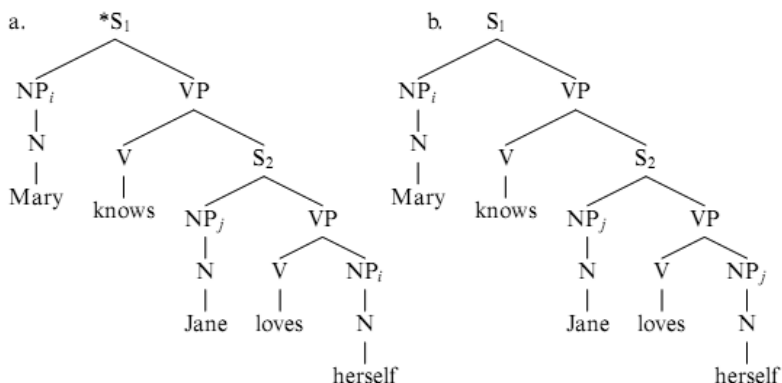
First, our examples of simple contrast between what is syntactically correct as oppose to incorrect. Take the sentence (H\*) “Herself loves Mary”, offered to John, and rejected by him. Why did he reject it? What is wrong with the sentence, that otherwise looks symmetrical to “Mary loves herself”, which is, let us suppose, OK for John? The answer is very simple for the professional linguist (Isaac 2008: 178), but less so for John’s general intelligence (and for mine as well, for that matter). The anaphor “Herself” has to be bound. In order for “Herself” to be bound, “Mary” should c-command it. However, the structure of the sentence, somewhat simplified, looks like this.



So, it is not the case that “Mary” c-commands “Herself”. In order to produce these answers, John’s internal parsing device should obtain the information that “Herself” is not syntactically connected to “Mary” in the right way. How can parser arrive at this? This can happen in two, or even three ways. First, something resembling the tree, call it “mental phrase marker” (I learned the term from David Pereplyotchik, so thanks go to him) can be either implemented in a non -explicit, non-representational way in the parser, so when the parser runs the parsing operation the result is that the sentence does not fit the mental phrase marker; the operation is either aborted, or some “red light” signal is emitted. Alternatively, the mental phrase marker is a full, explicit representation, and the parser “draws” it, in the way we did it here. Finally, and least probably, some items in the marker are merely implicit, others explicit.

Let us go along with the ordinarists and grant them the hypothesis that the mental phrase marker is implemented in an implicit way, so that there is no explicit representation in John's parser isomorphic to the tree we draw. If this is the case, what is the output available to the general intelligence?

We now pass to co-reference. Remember the sentence "(M) Mary knows that Jane loves herself". The linguist has put to the naïve subject John two questions, first, Does Mary know that Jane loves her, Mary, and the second "So, whom does Jane love?". The first question has been answered in the negative, and the second prompted the answer one would expect, namely that she loves herself, Jane. In order to produce these answers, John's internal parsing device should obtain the information that "herself" is in a right and complicated way connected to Jane, and not to Marry. It should have had at its disposition something corresponding to our two trees.



The crucial difference between them is the following: In the monster tree on the left-hand side, the first NP is co-indexed with the second NP containing the anaphor *herself*, which it also c-commands, and therefore binds, but the minimal clause containing the second NP is wrongly chosen: it is not the first, but the second one,  $S_2$ . So, whereas the conditions for binding are fulfilled, the locality condition is violated. (I am retelling the explanation from the textbook). The second tree fulfills both conditions: "Mary" binds "herself" (co-indexing and c-command are satisfied), and they occur together in the minimal clause, which is  $S_1$  itself. So, everything is legal. How can parser arrive at this? Again, this can happen in two to three ways, non explicit, non-representational way fully representational or mixed. The analogous story can be told for "They want to be teacher".

## 2.2. The intuition-core

Suppose we all agree that competence involves at least embodied and non-represented rules, and operates according to them. We also agree



that the immediate production is a datum. Michael and I went through several forms which the contrast between his ordinarism and my competentinalism might take, concentrating upon stages 2 and 3 of each view. First, the immediate answer of the competence, stage 2, and its origin, the tentative production of the sentence at stage 1. What does the immediate answer consist in?

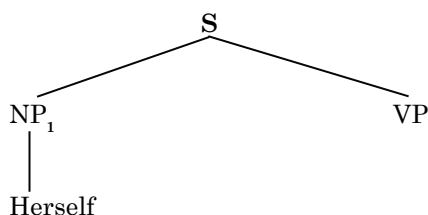
Dunja Jutronić mentioned once the extreme possibility that there is almost no answer at all, that competence just reiterates the sentence proposed, say “They want to be teacher.” That won’t do, since this is no new datum at all.

The second possibility has been put forward in Devitt’s answer to my criticism, in the context of discussing the question whether person’s answer to the linguist’s question is the datum:

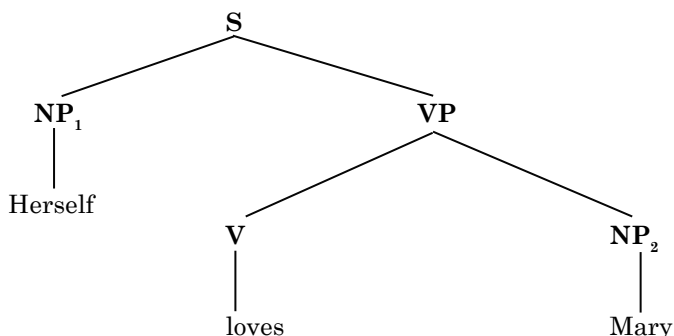
Her answer is not /i.e. the datum-NM/; it is part of the central-processor reflection. The datum is the experience that the answer is about. (2006c: 594 fn.22)

The experience, as made clear by the context, is “the experience of simulating the behavior” (2006: 594, the body of the text), i.e. the neural-verbal behavior of producing or trying to produce the target string.

This is hard to believe. First, subpersonal experiences hopefully don’t have qualitative character, so the experience of producing the string is just the very producing. Suppose that John’s competence or parser thus produces the marked string for “M”, and the CP takes this producing as its datum; it is almost like John’s CP watching the competence-parser producing the whole string. We may assume that upon receiving the word “herself” the competence looks in the dictionary and finds out that “herself” is a kind of word that can play the role in a noun-phrase. It hypothesizes the following simple structure:

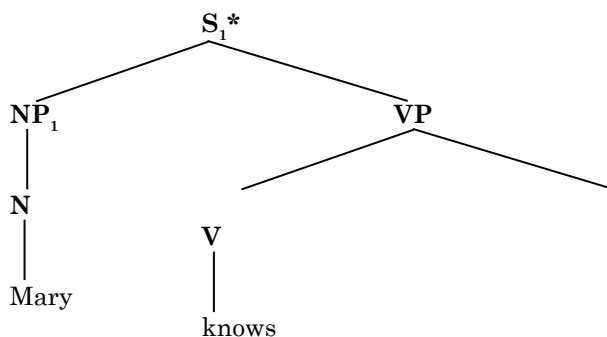


John’s CP is watching attentively, we presume on Devitt’s behalf. Next come “loves” and “Mary” so the competence happily merges them:



But what kind of information can this give to the CP? Linguistic rules are embodied in the competence, and not in the CP, since it is not a dedicated linguistic processor, but an all-purpose machine. So, mere following the toils of competence does not tell CP much. It has no idea about c-command, and can neither figure out that “Mary” should c-command the rest, nor that it does not do so in the tree. Only the competence has, or rather consists of procedural rules, so only it can decide whether the target string is acceptable. Devitt’s hypothesis that “The datum is the experience that the answer is about.” doesn’t tell us how CP could possibly figure out the verdict.

Things stand even worse with the second example. Assume for simplicity sake, and leaving the technical issues aside, that the parser just follows the order in which it is receiving input, and does not jump much ahead (is not a top-down parser). So, upon receiving “Mary” and “knows” and consulting lexicon, it disposes with the following “conjecture”:



What does this tell the CP? How much do you have to know to have even an inkling that a second sentence is expected on the right-hand side? And that an anaphora occurring in that sentence has to follow the rules of c-command? The two finished trees that we copied from the textbook are even more baffling for an ordinary CP. The second option is a non-starter, unless the ordinarist assumes that all people are the linguistic equals of Chomsky.

Let me introduce the third option, the “minimal signal” option, by wondering whether Michael perhaps means something else by his “datum”. Maybe, the competence ends up producing only correct strings; if the proposed string is ill-formed it stops, thereby rejecting it, like the “save” function in my word program, that just stops if the file name I want to save contains an illicit letter, thus letting me know that the title is not a correct string of letters. Here is a passage from Devitt that points in this direction; he is talking about “the normal competent speaker”:

If the datum shows that she would have no problem producing or understanding the expression, she is likely to deem it grammatical. If the datum shows that she has a problem, she will diagnose the problem in light of her background theories, linguistic and others, perhaps judging the expression ungrammatical, perhaps judging it grammatical but infelicitous or whatever. Often these judgments will be immediate and unreflective enough to count as intuitions. Even when they do count, they are still laden with such background theory as she acquired in getting her concept of grammaticality. (2006a: 109–10)

On this reading, at the end of its attempt to process the input string the competence signals “No problem” if the string is acceptable, or “I have a problem”, if the string is not acceptable (or, even simpler, its (re-)producing the string is the signal translatable as Yes, its having a problem and perhaps aborting the production the signal translatable as No.) So, just by following its toils to the bitter end, the CP can come to know its implicit verdict, in our example the negative one. The third option is just a development of this later alternative: that the answer is just a Yes/No signal in the neural code, translatable by the central processor. To make it more vivid, we can liken it to the red light on the crossing; the pedestrian interprets it as No, and stops if prudent. The ordinarist’s view is that this is very little, my view is that this is the core information, the real content of the intuition to be produced at the end of the day.

So, we do have here the voice of competence (albeit a discreet one, liable to be silenced by interferences) and it looks like intuition and feels like intuition. Devitt says explicitly that “(s)omeone who has the relevant competence has ready access to a great deal of data that are to be explained. She does not have to go out and look for data because her competence produces them.” (2006: 105) And the whole motivation of Devitt’s project was to get rid of immediate access to (deliverance of) competence. Finally, most of the work is done by the answer, our stage no.2, the agreed voice of competence. This happens in two ways. On the one hand, in many cases the stage no.3, empirical testing, does not change the verdict and adds nothing to it. In all these cases, it is the answer (no2) that is in its content identical to intuition (no.4), so its content just is intuition-content. On the other hand, when the verdict is somewhat modified at stage no.3., for instance when it is hedged (“I would never say this, but my kids say it all the time, so I guess it’s

OK”), the primary intuitional work is really done by the original verdict (stage no.2).

Consider again Devitt’s high demands on the naïve subject. Ann is supposed already to deploy her “folk linguistic concept of grammaticality” to appreciate the connection between this grammaticality and competence in the language. And she knows that she is a competent speaker and so uses herself as a guide to what the competent speaker would do. So she asks herself whether this expression is something she would say and what she would make of it if someone else said it. Her answer is the datum.” Does the linguist really need all this from Ann? Devitt himself kindly suggested in correspondence the negative answer. The basic data for the linguist are that Ann would say such-and-such and would mean such-and-such by the given expression. The rest is sociological theory. Remember, when a Chomskyan speaks about one’s language, he means one’s idiolect. Ann’s opinion whether she is a typical representative of a community, whether her English is good English are beside the point. Does Ann have to think that she is competent in the very language (idiolect) she is speaking? The very question sounds ludicrous. Why is this relevant? Because stating the judgment about a sentence in one’s idiolect requires *far less theory, if any at all, than reflection about social usage.*

The fourth possibility is that the answer is a relatively articulate verdict, of the kind “No, this is not a well-formed string”, again formulated in a neural code. I suppose that Michael rejects it, I would leave it as an open possibility.

To reiterate, I find first the option empty, and the second option, that the datum is experience of simulation itself extremely implausible if taken literally. I tend to agree with the minimal signal option of Yes/No signals, green/red lights in a neural code (and with its subspecies, the more tolerant alternative reading of Michael’s “the experience that the answer is about”, according to which the successful experience is a green light for the string, the aborted simulation a red light.) It has the advantage of offering a distinct job to competence, and a distinct job to CP, it goes well with poverty of intuitions, their tendency to reduce to Yes/No final verdicts, and it is very parsimonious. However, I have nothing against the articulate signal option, and would be happy if psychologists confirmed it, since this would make more space for competence and help its voice to be heard more easily. End of stage 2.

### 2.3. *From signal to intuition*

This brings us to the work of CP, our stage 3. Here, two opposite possibilities loom large. Either the empirical theorizing is narrowly linguistic or it is wide, holistic, and could involve language-external, for instance social, affective and other considerations.

So, to summarize, the ordinarist can propose either the wider or the narrower task(s) for the CP. If the task is wider, it is not specifically

linguistic; the linguistic job is done by the competence, and the core of linguistic intuition is really the decision of the competence, it's red or green light. If the task is narrow, it is just translating the very same datum of the competence. Either way, it is the competence that provides the essential core of the resulting intuition: either its linguistic core, on the wide picture, or its total content, on the narrow one, as predicted by The Moderate-Voice-of-Competence view.

How does intuition-capacity develop? We know very little about the improvement of intuition capacity. Some researchers mention that illiterate people did not even understand the questions they asked them: the people just have no idea of what it could be meant by asking whether one would say certain things. Robert Matthews provides a piece of evidence Devitt reports

This point is nicely illustrated by the following report: "As a graduate student I spent a summer in the Pyrenees (Andorra, Perpignon, etc.) doing field research on the phonology of various dialects of Catalan. Many of our native informants were illiterate peasants. I was forcefully struck how difficult it was to elicit linguistic judgments from them regarding their language, which of course they spoke perfectly well. Just getting the plurals of certain nouns was tough. These folks seemed to be very hard of hearing when it came to hearing the voice of competence! Their difficulty, it seemed, was that their native language was largely transparent to them—they had never thought of it as an object for observation and hence were largely unable to form even the most rudimentary judgments about its character. Catalan speakers with only a modicum of grade school education, by contrast, were good informants, presumably because they had learned through their grammar lessons to think of language as an object with various properties, even if they had no sophisticated knowledge of what those properties might be, theoretically speaking." (Bob Matthews, in correspondence). (2006a: 109 n)

On one reading, the one I prefer, the informants can't distinguish asking about linguistic correctness from asking about pragmatic appropriateness. Maybe, that is, they had difficulties understanding Matthews' questions: is this gentleman asking whether I would say this under some imaginable circumstance (e.g. when drunk, or joking), or whether it would be an appropriate statement to make, or something else? It is hard to believe that in their normal life they do accept systematically ill-formed utterances: in a village, a person with linguistic deficit might be severely ridiculed, and the ridicule seems to speak in favor of other villagers having very definite intuitions to the effect that something is badly wrong with the poor person's way of talking.

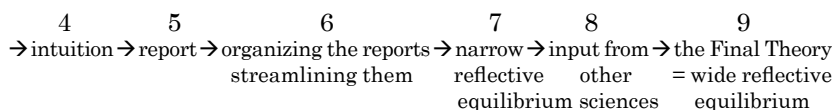
Let me conclude with a few short remarks about the list of problems in waiting, connected to the notion of basic competence, as developed by Chomsky. How much of its power is due to the innate structure, how much to empirical learning? Let me, with the majority, call the descriptive-explanatory view that stresses empirical learning "empiricism", and reserve the term "rationalism" for the opposite descriptive-explanatory view, that experience plays no essential role in the process, beyond mere triggering or prompting. (I shall later be calling

the corresponding normative views, concerning justification “aposteriorism” and “apriorism” to avoid the confusion between the normative and the descriptive-explanatory.) The two dynamic processes, relevant here, are first, acquiring the language, and second, improving one’s intuition capacity, and they both constitute an important area of research. Both seem to involve straightforward learning. Even on the most mainstream Chomskyan view, which Michael of course rejects, the child’s language module needs external information to fix the parameters characterizing the language of the surroundings. In later stages, the child will learn more and more sophisticated facts about the language that is going to become her mother-tongue. Does this imply (descriptive) empiricism about intuition? Only a bit of it: the scaffolding is given, but some information is learned from surroundings. And it might be interesting to discern the fine structure...

In my paper “Intuitions: the discreet voice of competence” (2006), I have argued that our innate endowment might explain at least the very origin of the basic intuition-capacity and the initial stages of the formation of our intuition-states with their contents, but that nativism should be restricted to the origin of the system and to the relatively initial stages of processing. Does the incontestable fact that people’s intuitions do develop, and that sophisticated, well-trained people have much richer intuitions speak in favor of empiricism? Again, only to some extent.

Let me conclude this part of the story by stressing the passion for explanation, omnipresent in the debates about linguistic intuitions. This passionate “explanationism” has been in the past characterizing the debate about all sorts of intuitions, in the work of classics, like Plato, Descartes, Kant, and the early twentieth-century philosophers like Russell and Husserl; it is a pity that present-day debate on intuitions in general is much less marked by it.

How do the stages of intuition-production fit with the theoretical work in linguistics? The main way a linguist arrives at her theory is nowadays by focusing upon speaker’s intuitions. We may add this development to our story about stages. Remember, the two last stages were the explicit intuition and the report. We now add three more stages. The first is putting intuitions together, and streamlining them (weeding out contradictions, mistakes and the like); it corresponds to building a narrow reflective equilibrium in the case of other kinds of intuitions. Then comes narrow grammatical-linguistic theorizing, and finally balancing the results with the input from other disciplines: psycholinguistics, neurology, possibly evolutionary biology, and science that could help. The result, the Final theory, would be something like a wide reflective equilibrium, encompassing intuitional and other data, input from other sciences, and linguist’s theorizing.



So much about intuition. We now turn to the other set of questions, the ones concerning the metaphysical status of linguistic entities.

### 3. *Validity of intuitions:*

#### *Production-and response-dependence*

##### 3.1. *SLEs are production-and-response-dependent*

The idea proposed and briefly defended above that intuition is just another window upon the linguistic competence suggests that language production, understanding and judging go together. It helps to make our view of language abilities more unified and systematic. Moreover, such a unified picture suggests also a metaphysical framework for locating language. Here, the question we want to answer is: where is language situated in the general order of things? In the mind, in the world, or somewhere in-between, straddling the divide. Call this question, following the usual usage in metaphysics, the location problem for language. Here is how it arises.

Commonsensically, the way the folk hears and sees the linguistic tokens seem to suggest that the language is in the outside world, it is “the spoken and written stuff”, with no narrower specification. It is what was later called E-language by Chomsky. But the folk wisdom has been put in doubt, since rather early times, at least since the second century AD. Sextus Empiricus in his treatise *Against the Grammarians* questions the existence of language: for him, the availability of sound, syllable, word and sentence are highly dubious. Sounds are questionable since we don’t have a clear principle of identification that would tell us, for instance, if a diphthong is one sound or two or whether a hardly audible “r” in a syllable is a sound or not (I, 117 ff.). Words and parts of the sentence, inherit the dubiousness; “if the aggregate of the parts of the sentence is conceived to be a sentence, then because the aggregation is nothing apart from the parts aggregated, just as distance is nothing apart from the objects which are distant, the sentence of which any parts shall be conceived will not be anything. And when the whole sentence is nothing, neither will any parts of it exist. I. 135, p. 81). The conclusion is that “neither does the sentence exist” (I, 137). Many contemporary linguists and philosophers of language would agree with Sextus: in their view, already phonology shows that there is a problem: the phonetics-phonology interface is complex, and what is heard depends heavily on context, and more importantly on expectations and habits. This internal, mental component has become even more central for the linguists with the advent of Chomskyan linguistics. The question now looms large: where should we locate the language?

Let me detail the problem a bit. Chomsky's insistence upon the importance of I-language has generated the temptation of radical internalizing motivated by the wish completely to get rid of the first, external component, that is seen as too contingent, "non-linguistic" and unsystematic to be worthy of being called "language" in the strict sense at all. "Language, as far as I can tell, is *all* construction" writes Jackendoff (1992: 104, emphasis by Jackendoff), echoing Sextus after almost two millennia. He also initiated the comparison between "hearing" the linguistic items and "seeing" visual illusions. The invidious comparison is later developed by G. Rey (2005) a fervent defender of the picture of language as merely internal, even "imagined" and also used by Isaac (2008: 24). Phonemes, their aggregates (morphemes, internal sentence-strings), syntactic trees and other "standard linguistic entities" (SLEs to use the abbreviation preferred by G. Rey) are in the head, and they are the whole of language. Language is I-language, E-language is out.

Another example is the blank screen metaphor offered by John Collins: E-language is like the blank screen onto which linguistic items are being projected by our mind. It sounds there is literally nothing out there. Take the book with complete works of Shakespeare, and take an empty book, just blank pages like movie screen; it can't be that in the case of his complete works there is nothing (written in a language) there, that we project the text upon blank pages. Collins has responded in the correspondence that he attacked only extreme objectivism, without really wanting to deny that there is E-language outside. But saying, as he does, that "the tokens are there alright, but the properties are simply projected onto them" bring the blank screen back in; the tokens don't have any linguistically interesting properties.

Philosophers are sometimes tempted to extremes: in this case to extreme internalization of matters linguistic. The temptation is to halve the reality of language. Commonsensically, the language is outside, it's the spoken and written stuff. Comes Chomskyan linguistics and detects another half, the internal one, which turns out to be quite important. This generates the radical internalizing temptation to get rid of the first half. But such a Sextan line has to be resisted, on the pain of linguistic nihilism. Commonsense entities should not be liquidated unless it is really necessary. Widespread illusions are not to be postulated unless this is absolutely unavoidable. Notice that utterances (speech, books, conversations) are (still) part of explanandum in linguistics, so E-language is (still) part of the linguistic enterprise. The E-language is needed first as crucial input: you need utterances that are acceptable or not. Note the importance of the gathering of data, and the fanatical attention to the details in the data that characterizes great linguists. The strong internalizing move tends to make nonsense out of this importance, and turns linguistics into a rather strange enterprise of non-accounting for anything at all.

Devitt prefers a strong externalizing move. Language is just a piece of non-mental reality, it's physical signs and body-involving practices,



and nothing mental at all. He contrasts linguistic reality to mental reality.

Second, the theory building proceeds by accounting for the acceptability of these E-language items. The impressive machinery of UG is consistently being proposed as an account of reactions of native speakers to utterances proposed or thought of (when the linguist is using herself as the linguistic guinea-pig) limited variability of data. The data cannot vary indefinitely and in a chaotic fashion. Especially if simplicity is such a prime ambition: if all of the grammar is to be captured in a few very simple and very general principles, then the data should be highly organized: all the data should be amenable to the principles. Why is the idea of parameter setting so impressive: because it detects a deep, very general uniformity, and constraints variation with an iron hand!

Finally, note the contrast between folk-superstition and folk view of language as being out there in the world. The first ascribes all sorts of crazy powers to presumed witches, which the women in question just don't have. Folk-views on language don't ascribe any crazy stuff to sentences. They just fail to add sophisticated stuff, that's hardly an error. So, we need a way to preserve the folk or commonsense view that utterances *are* language.

On the other hand, the path back to the Eden of commonsensical firm objectivism is also closed, since too much of linguistic material is clearly located strictly in human cognitive apparatus. So, why not go part of the way with the folk, give unto complete works of Shakespeare what belongs unto them, and give to the mind what belongs unto it. The answer I favor is therefore that the two halves, the internal and the external, E-language and I-language, have to be kept together. This is compatible with a wide range of views, from eliminativism (linguistic items are only projected)<sup>1</sup> and quasi-eliminativism about them <sup>2</sup>, to various views that stress speaker's intention as basic for linguistic product(ion)s to the classical Chomskyan stance according to which language is essentially a mental thing, and only accidentally has an external manifestation. So, how do we choose? I think that the best way is the most direct one: language is relational and response-dependent. But the answer needs a lot of developing and defending: pointing out that it is in the middle between the two extremes would have been able to recommend it in times of Aristotle, when the virtue was considered to lay in the middle.

In brief, linguistic reality of a language, say French, is pretty much determined by the psychological reality of its speakers. The mental dispositions and states of native French speakers, determine in the last instance, the structure (syntactic, semantic and pragmatic) or our language. Capacity of recognizing the correct forms, crucial for intuitions, is part of this constitutive structure. A sound-string is a correct sentence of a language partly because a competent language speaker would recognize it as such. It is obvious that *p* is a correct sentence of

the language iff a normal thinker would find  $p$  clear and compelling. In this sense, the validity of intuitions is ultimately explained by the efficiency of the wider collective psychological make-up of which they are part. The order of determination seems to go from the mind to the world. Language is in-between the mind and the world, straddling the divide. Let us look at more detail.

Start with the level of phonology. Why is a certain stream of sounds, or a series of inscriptions a token of “M”? Because John would have intuition that it is, and hear it and read it as a token of “M”. To pass further to graphemes, why are the following two very differently looking strings in fact tokens of the same Croatian sentence?

- (L) Svaki od dva kandidata očekivao je da će pobijediti onog drugog.  
 (C) Сваки од два кандидата очекивао је да ће побиједити оног другог.

Because a Croatian reader knowledgeable of the Cyrillic script would see them so, and, if need arose would write his thought in the form of (C). I just did it on my computer, and I count as a Croat with a command of Cyrillic.

So, the linguistic properties of the sentence seem to be dependent on the production and on the responses of the speaker in charge. Let me then introduce a new term, with a play of words to boot. Call this hypothetical dependence “*production- and response-dependence*”, *PR-dependence for short*. You might take the “PR” as a pun, reminding one of “public relations”, which is not bad: a more psychologically minded reader might read the term strictly as referring to dependence on individual speaker-hearer, a more sociologically minded one as pointing to further dependencies. So, a simple relational answer to our metaphysical location problem would be that linguistic entities are production- and -response dependent. In the form of a slogan, *SLEs are PR-dependent*. (Thanks go to Michael who in discussion warned me that pure response-dependence is very, very implausible; he would not accept the PR-dependence either.

Let me explain the slogan and argue for it in more detail. The slogan claims the following about any given SLE:

SLE characterizes a given linguistic (external) token  $t$  of L iff a normal speaker-hearer of L would accept  $t$  and would produce  $t$  if suitably prompted.

And, the left-hand side determines the right hand one: SLE is a property of a given linguistic (external) token  $t$  because the normal speaker of L would do as specified. With the advent of Chomskian linguistics it has become a commonplace that competence in a way *dictates* what counts as the correct, well-formed sentence of a given language. Let us pass to illustrations and explanations.

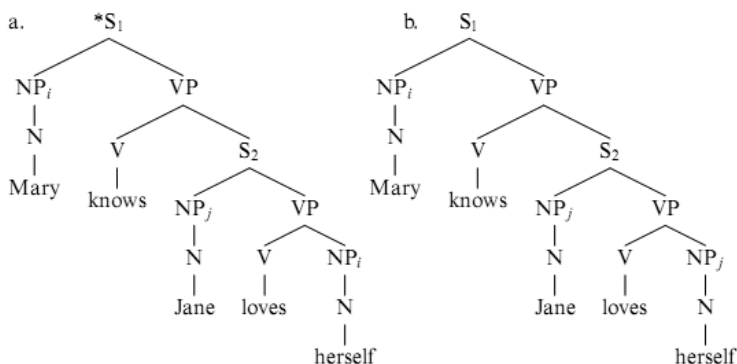
Start with the lowest level. The commonalities between different tokens of the same sentence like our (L) and (C), but also of tokens variously articulated, are recognizable only through the detour through the

response of the hearer, call it response-detour. The commonalities are also *explainable* only in terms of the response-detour. Therefore, by inference to the best explanation, the fundamental phonological (-graphematic) properties of the strings are not mind-independent properties. An external token string has linguistic properties because it would be interpreted as having them by the normal language-hearer, and would be produced by a process that would form it respecting the nature of these properties. So, they are best seen as PR-dependent properties

The same is valid for fundamental grammatical (-semantic) properties of the strings, bringing in our discussion of intuitions. We focus here upon the first family, the syntactic one.

In researching a given language the linguist has to do with speaker-hearer, the person who not only produces a string, but also reacts to the string itself, so that her reaction, in the form of her intuition and report, offer a glimpse into the structure of her I-language. Return for the moment to the sentence

(M) Mary knows that Jane loves herself.  
and to our two trees.



Now, why does the sentence “M” say that Jane loves herself, Jane, and that Mary knows this? Why not that Jane loves Mary; that might have been a happier situation. Well because John, the speaker-hearer of the language, has intuitions that can be systematized by the two trees, and has, presumably, something in his cognitive apparatus that is roughly isomorphic to them (either explicit representations, or a sequence of processing moves). If he heard “M” systematically in the “Jane-loves-Mary-who-knows-about-it” way, and produced it to indicate this, then the sentence would have had a different (a linguistically “illegal”) syntactic structure represented by the left-hand monster tree. Moreover, John would himself produce “M” to mean that Jane loves herself, Jane, and would thus testify to the fidelity of the “legal” analysis. The two combine in simple cases. (The exceptions and complications concern the “impure” intuition we talked about earlier, for instance our hero Ina who allows ungrammatical strings out of her love for the non-na-

tive partner, but does not produce them herself. The other, and indeed famous example is the opposite one, studied by Labov: the snobbish reaction of not officially allowing utterance-tokens of the kind one produces oneself).

So, the property of the given utterance of “M” of being a token of a sentence with the c-command pattern that makes it into a correct piece of John’s language, looks like a relational property, that the token has in virtue of playing a (potential or actual) role in awakening John’s responses, including the intuitional ones, and of being potentially produced by John’s linguistic apparatus as a correct item. So, in general it seems that the relevant syntactic properties are relational properties, that the physical type sentences possess in virtue of being related to psycho-grammar (in terms of being produced/-able and being recognizable/parsable/ as such), i.e. PR-dependent properties. The production can be taken in one of the two ways, and I will remain officially neutral between them. For the representationalists about SLEs, respecting a mental phrase marker is to be taken literally: when John is producing the sentence “M”, his mental apparatus constructs the marker and then organizes the sentence in accordance to the representation produced (merging and moving its elements with one eye on the represented paradigm). A non-representationalist will settle for less: the apparatus functions as if it is implementing the phrase marker, it follows the rules without representing them, the way planet’s follow Kepler’s laws, blindly but reliably. In both cases the best choice is our slogan: SLEs are PR-dependent.

But isn’t the tree still merely projected onto the sentence-token by John’s mind, a strict internalist might ask. In answer, we should think of degrees of projection versus guidance. The first and clearest is total projection upon a blank screen, like the movie of tv-screen. The next is moderate projection like in the Rorschach test, where the external shapes play an auxiliary suggesting role. The third is a mix, combining minimal projection with moderate guidance. Take projecting depth or perspective into a realistic-style painting of an imaginary landscape. The painter has done most of the work, her hand guides the eye of the observer; still, an observer with no experience with paintings or photographs will have trouble noticing depth. Consider how the three cases reflect on the study of the relevant phenomena. First, nobody studies blank screens in order to understand various movies, but, second, psychiatrists propose better or worse test, without normally trying to account for every point in the drawing within a systematic theory. And finally, concerning the recovering of depth-information, the typical art historian pays maximal attention to almost every feature of the realistic-style painting. The painting shows with great exactness what *the painter had in mind*. Linguists come *closest to the third case*, with endless, patient search for examples and counterexamples, pairs of very similarly looking written sentences, one of which is acceptable

and other not. So, the right mix for language is *minimal projection and maximal guidance*. The corpus shows with great exactness what *the speaker had in mind*, sub-personally, of course. And indeed, a volume of prose or drama in a given language hardly reminds one of blank screen; a post-modernist might see in it a kind of Rorschach, but the scientific linguistic attention to detail points to the opposite end, the minimal projection, exactly as the PR-dependence picture would predict.

Let me conclude by mentioning several interesting issues related to the narrow response-dependence, i.e. dependence on the hearer. Suppose that I want to say that Mary is aware of Jane's love for her, but, having previously had a nice portion of my favorite brandy, I produce a token of "Mary knows that Jane loves herself". "You claim she knows that Jane is narcissistic?" a colleague asks. Well, this is what I have said, and here it is the hearer, the colleague, and her response, that decides. The same goes for meaning, in particular public meaning vs. speaker's meaning. The speaker thinks that "promiscuous" is a very posh word for "promising" and says: "I like promiscuous candidates." He meant "I like promising candidates", since this is how he composes (produces) and hears (responds to) his sentence. Unfortunately, the public meaning is the other, slightly awkward one. Why? Well, my dear speaker, because most people would hear your sentence as talking about candidates who change their partners a bit too often (for some standards). Looks like the public meaning is pretty much narrowly response-dependent, defined by how people would decode the utterance. Now, if (and I say *if*) the public meaning has primacy, then the hearer is privileged in relation to the speaker, and we have strict, narrow response-dependence.

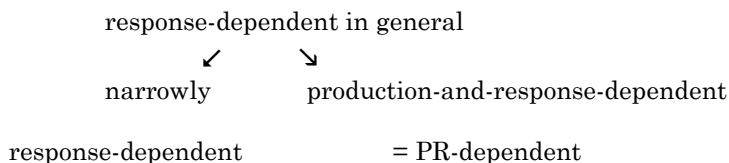
On the other hand, in the case of ambiguity, as Devitt has warned me in a pleasant discussion in the IUC in Dubrovnik, the speaker seems to have the upper hand. If I say, "Visiting relatives can be boring", meaning the ritual of visiting them, and you take it the other way around, it is me who is the authority. So, dependence on production remains, along with the one on response. SLEs are PR-dependent.

### 3.2. *Sketching the big picture: Intuitions and the maximalist response-dependence view*

How does the view just sketched, fit with the rest of our world picture? Are there other PR-dependent properties? What about the moral properties like *right*, *just* and *fair*? Many people swear these are fully objective, others see them as "social constructions" or projections. But there is a middle way. One can say that an action was just if it would strike an impartial observer as being such (or, impartial observers would agree to this effect, under one or other guise, including veil of ignorance and similar devices). Kantians would insist on the production: the action was right only if it was produced with the right intention, the one of doing what is morally required by the situation. If we combine the two, we get the PR-dependence:

An action *a* is morally right iff it has been produced with the intention of doing what is right and is perceived by impartial observers as right.

This line has not been explored, and I rest content with just mentioning it here. (I got the idea of the moral-linguistic parallel in a conversation with Nenad Smokrović, so thanks go to him). Obviously, the language, if to some extent mind-dependent, is certainly not alone in being such. For pure response-dependence take secondary qualities like *bitter* and *red*. They are hardly production-dependent, but it is arguable that they are response-dependent, and I have argued for the view in several papers. We may introduce the term “generalized response-dependence”, to cover items like SLEs that are both production-and-response-dependent and those that are response-dependent in strict sense. This would yield the following division:



Now, all the examples (*bitter*, *red*, SLEs) are cognitive, as opposed to moral-practical. We might add other cognitive examples of the same high level, for instance being obvious; it is a property typically accompanying intuition-contents. When is it obvious that something is the case, e.g. that our ( $M^*$ ) is an incorrect sentence:

It is obvious that *p* iff a normal thinker would find *p* irresistible, clear and compelling.

And, the left-hand side determines the right hand one: it is obvious that ( $M^*$ ) is incorrect because the normal speaker of John’s language would find it compelling that it is so. Our next non-practical example are expressive properties, like e.g. *happy-looking*. A person is happy looking because it looks happy to her surrounding, more officially because it would look happy to the relevant group of normal observers. What about individual emotional qualities like *frightful* or *sad* as predicated to objects and situations? They might be response-dependent as well: a scene is frightful because it would frighten the normal observer. Our last non-practical example is much more philosophical, aesthetic properties like *good-looking*, *beautiful* and *ugly*. Since Hume’s times they have been characterized by many philosophers in response-dependence terms.

Passing now to the minefield of other practical candidates for response-dependence beside the moral qualities, let me start with the least problematic candidates. Take individual attraction and repulsion, with qualities like *attractive*, *disgusting*, *sexy* and the like; they are probably the least problematic candidates for a response-dependence

account. Social properties like *inviting* and *offensive* follow suit. Like in the case of language I am in favor of a response-dependantist view. Similarly with the “significance” properties, like *meaningful* or *meaningless*, particularly for items like a meaningful life.

This is then the first sketch of the big picture: many, perhaps all humanly interesting properties are response-dependent in the very general sense indicated (involving production-dependence besides the narrow dependence on the response of the hearer); language is here joined by its peers within a huge family of properties.

#### 4. *Conclusion: Explaining intuitions*

Let me recapitulate. In the first part I have defended a moderate “voice-of-competence” view, the view that there are intuitions-dispositions and judgments, which form a distinct group of phenomena, and there is the intuition-capacity, the capacity to use our linguistic competencies in an off-line fashion. It is the voice of competence, most often a discreet one. Intuitional data are thus the minimal “products” of tentative linguistic production and primarily not their opinions about the data. The data involve no theory and very little proto-theory. Although there might be admixtures of guesswork in the conscious production of data, these are routinely weaned out by linguists. In contrast, ordinarism denies that seemingly linguistic judgments form a significant epistemic kind, intuitions, and that there is a distinct capacity producing them. I agree with ordinarist about referentialism: intuitions are concerned with their external objects, the domain of items and facts. I agree about the importance of explanationism in contrast to quietism: items traditionally described as intuitions require an explanation of having and reliability, if possible a causal one. However, I have defended against ordinarists the competence-centered option, that is also referentialist and explanationist, and proposed a sketch of explanation, featuring the hypothetical stages of intuition-generation and the less hypothetical ones of reflection about intuitions.

The second part develops a very rough sketch of what in the reality validates our linguistic intuitions. For this, it turns briefly to the metaphysics of language, placing it in-between two extremes, the purely psychological Chomskyan one and the purely extra-mental, the Devitt’s one. With the advent of Chomskyan linguistics the old question of where language is situated in the general order of things has been partly answered by stressing the role of the mind. The order of determination seems to go from the mind to the world. The view defended here agrees but adds and stresses that language is in-between the mind and the world, straddling the divide, and that a sound-string is a correct sentence of a language because a competent language speaker would recognize it as such. To put it in a more technical sounding terminology, language is production- and response-dependent, and being response-dependent goes together with color, moral, epistemic and aesthetic

value and with emotional properties. The two parts offer a sketch of a relatively encompassing view of language and linguistic intuitions, combining epistemological and metaphysical topics.

## References

- Collins, J. 2008. *Chomsky: a guide for the perplexed*. London: Continuum.
- Devitt, M. 2005. "There Is No *A priori*." In E. Sosa and M. Steup (eds.). *Contemporary Debates in Epistemology*. Oxford: Blackwell Publishers.
- Devitt, M. 2006a. *The Ignorance of Language*. Oxford: Clarendon Press.
- Devitt, M. 2006b. "Intuitions in Linguistics." *The British Journal for the Philosophy of Science* 57: 481–513.
- Devitt, M. 2006c. "Defending Ignorance of Language, Responses to Dubrovnik Papers." *Croatian Journal of Philosophy* 6: 571–605.
- Devitt, M. 2012. "The Role of Intuitions." In G. Russell and D. Graff Fara (eds.). *The Routledge Companion to Philosophy of Language*. London: Routledge: 554–563.
- Ewing, A. C. 1971. *Fundamental Questions of Philosophy*. London: Routledge.
- Isac, D. and Reiss, Ch. 2008. *I-Language an Introduction to Linguistics as Cognitive Science*. Oxford: Oxford University Press.
- Jackendoff, R. 1992. *Languages of the Mind*. Cambridge: MIT Press.
- Mišćević, N. 2006. "Intuitions: The discrete voice of competence." *Croatian Journal of Philosophy* 6: 523–548.
- Moore, G. E. 1991. *The Elements of Ethics*. Edited and with an introduction by Tom Regan. Philadelphia: Temple University Press.
- Rey, G. 2005. "Mind, Intentionality and Inexistence." *Croatian Journal of Philosophy* 5: 389–415.
- Sextus Empiricus. 1949. *Against The Professors, Book One, Against the Grammarians*. Cambridge: Loeb.